Teachers' Self-Efficacy: progressing qualitative analysis.

Abstract

Teacher self-efficacy has predominantly been explored using quantitative instruments such as Likert scales-based questionnaires. Several researchers have questioned these methods, suggesting they offer only a limited view of the concept. This paper considers their claim by exploring the self-efficacy of UK secondary science teachers participating in a two-year professional development programme using both traditional quantitative scales and qualitative methods, including interviews and lesson observations. The findings support the suggestion that traditional quantitative scales do not fully capture teacher self-efficacy and highlight inconsistencies between self-efficacy assessments through the different research approaches. We argue that to achieve a more complete and comprehensive picture of teacher self-efficacy, it is essential that traditional quantitative approaches are better triangulated and integrated with other sources of data, in particular lesson observations. We offer an emerging approach of how qualitative data sources might be used to develop this comprehensive picture.

Key words: teacher self-efficacy; qualitative methods; outdoor science education; teacher professional development.

Introduction

Self-efficacy relating to teaching ability is undeniably linked to successful teaching (Bandura 1997). By evaluating teacher self-efficacy, important information is provided to researchers, policy makers and teacher educators, as to, for example, the success and future design of professional development programmes (Wyatt 2012) and predicting teacher commitment to the profession (Chesnut and Burley 2015). However, the assessment of teacher self-efficacy poses multiple challenges that require resolutions (Henson 2002; Goddard, Hoy & Woolfolk-Hoy 2004). For example, over the past 20 years the majority of teacher self-efficacy research has used quantitative tools (for example, Likert scales-based questionnaires) (Klassen et al. 2011) which Wheatley (2002) asserts, due to the lack of qualitative research, has resulted in a narrow conceptualisation. Furthermore, the reliability and validity of current quantitative instruments, together with teachers' self-report, have been called into question (Pruski et al. 2013). Wyatt (2012) recommended that if research is to be of real use to teachers and teacher educators, the development of in-depth qualitative self-efficacy methods are crucial.

Hence, the rationale for this paper is two-fold. First, in response to the lack of qualitative self-efficacy studies, we present emergent qualitative methods for the collection and analysis of data for teacher self-efficacy. Where self-efficacy has been explored qualitatively, few researchers have used a range of data sources and set out an explicit framework for analysis, enabling transparency and presenting an opportunity to begin to develop a systematic qualitative approach. This paper is unique in that we collected and incorporated lesson observation data into our analytical framework, a data source rarely utilised in self-efficacy studies (Wyatt 2012). The second rationale for the paper is to confirm, and then demonstrate, the need for more than just traditional Likert scale approaches for self-efficacy research. To this end we collect, compare and examine both qualitative and quantitative data to assess the participant teachers' self-efficacy. This article ultimately addresses two questions:

- How might qualitative methods be used to gain a more comprehensive picture of teacher self-efficacy?
- What, if any, variation is there between teacher self-efficacy assessment when qualitative and quantitative methods are applied?

The context for this research was the 'Thinking Beyond the Classroom' programme (see, www.pstt.org.uk), which provided professional development for secondary science teachers in England to teach science outside the classroom. As one of the programme's objectives was the development of teacher expertise and teacher self-efficacy for outdoor science teaching, the context was conducive for this study.

Understanding Self-Efficacy

Self-efficacy resides in a complex psychological landscape. Bandura (1997) argues that the construct occupies a pivotal role, acting upon other classes of determinants – including beliefs, attitudes and motivation. From a social cognitive perspective, teacher self-efficacy, a sub-category of self-efficacy, has been defined as a teacher's judgement of their ability to influence student outcomes (Klassen et al. 2011); or more specifically, it is 'a teacher's individual beliefs in their capabilities to perform a specific teaching task at a specified level of quality in a specified situation' (Dellinger et al. 2008, 752).

The concept is considered useful as by identifying both a person's 'perceived self-efficacy' and their outcome expectancy of a particular task, subsequent behaviours can be predicted (Bandura 1997). For example, a teacher with high teacher self-efficacy and high outcome expectancy is thought to behave in a productive manner and gain personal satisfaction from the endeavour. In contrast, a teacher with low teacher self-efficacy and low outcome expectancy is predicted to exhibit behaviours displaying resignation, and an attitude of apathy (Bandura 1997).

Increasingly, research reports specific behaviours, or teaching traits, aligned with particular judgements or levels of teacher self-efficacy. Teachers categorised to have 'high teacher self-efficacy' have been observed to demonstrate the following attributes and/or pedagogical practices, including: good subject knowledge (Riggs 1995; Ross 1998; Muijs & Reynolds 2002; Cantrell, Young & Moore 2003); high levels of planning, organization and enthusiasm (Allinder 1994; Muijs & Reynolds 2002); and a willingness to experiment with new teaching methods, persisting even if initial implementation is flawed (Haney, Czerniak & Lumpe 1996; Cousins & Walker 2000). Furthermore, teachers who are reported to have high teacher self-efficacy use a variety of teaching methods (Cantrell & Callaway 2008), are less reliant on curriculum guides and emphasise cross-curriculum links during teaching (De Laat & Watters 1995). Such teachers often emphasise problem-solving and logical thinking through 'real life'

examples valuing the 'unique' opportunities for student autonomy rather than viewing real life example solely as fun (Czerniak & Schriver 1994; De Laat & Watters 1995; Glackin 2016).

Finally, teachers with high teacher self-efficacy were more likely to use difficult to manage methods, such as inquiry, small group work and cognitively challenging questions, and less likely to adopt didactic practices (Ashton & Webb 1986; Muijs & Reynolds 2002; Chacon 2005). They were noted to have relatively few controlling beliefs about classroom behaviour (Woolfolk, Rosoff & Hoy 1990; Chacon 2005); and, were relatively uncritical of students who made errors and persisted in supporting students who struggled (Gibson & Dembo 1984; Ashton & Webb 1986).

Whilst contexts and methods varied between the studies from which these findings were drawn, reported behaviours for particular levels of teacher self-efficacy were broadly consistent. The significance of this shared list of attributes and practices for teacher self-efficacy is that it offers a rich description of features that are identifiable across teachers' practice and through interviews.

Assessing Self-Efficacy

Investigation of teacher self-efficacy over the past three decades has been dominated by quantitative methodologies. Klassen and colleagues (2011) reported that 76.7 per cent (n=167) of the studies of teacher self-efficacy they reviewed (1998-2007) were conducted using quantitative methods. In general, such methods have relied on teacher self-report, involving the development and use of validated surveys and standardised protocols. The majority of such instruments use a two-scale dimension reflecting Bandura's assertion that self-efficacy is the result of personal factors and the environment. The Science Teacher Efficacy Belief Instrument (STEBI), for example, designed by Riggs and Enochs (1990) and more recently revalidated by Bleicher (2004), consists of 13 statements relating to the perceived self-efficacy dimension and 12 relating to the outcome expectancy dimension (Riggs & Enochs 1990). These are exemplified by, *I am continually finding better ways to teach science,* and, *the inadequacy of a student's science background can be overcome by good teaching,* respectively (Riggs & Enochs 1990). Teachers are asked to indicate the degree to which they agree or disagree with the statements on a five-point scale, running from *strongly agree to strongly disagree.*

Instruments for measuring teacher self-efficacy have evolved over several decades. For example, the STEBI, arose from earlier instruments such as RAND (developed by researchers at the RAND corporation; Armor et al. (1976)), Teacher Efficacy Scale (Gibson & Dembo 1984) and Ashton, Buhr and Cocker's (1984) vignettes. In turn, the STEBI has informed such measures as the STEBI-CHEM that measures teacher's self-efficacy in teaching chemistry (Rubeck & Enochs 1991). More recently, Bandura (2006) has set out guidance for constructing self-efficacy scales. The instrument's continual development is often prompted by the desire for increased context/subject specificity (Tschannen-Moran & Woolfolk Hoy 2001).

More generally, quantitative approaches dominate psychology-related research. This, Gardner (1996) proposes, is due to their ease-of-use, transferability and the fact that multiple questions can be used to measure a construct – potentially increasing reliability. In terms of teacher self-efficacy *per se*, we would argue that the extensive use of quantitative methods is the result of the theoretical framework underpinning the construct. Explained further, self-efficacy is rooted in an epistemology that behaviour is determined by both a person's cognitive processing capacity and the environment within which they exist, each thought to influence one another, determining the resulting behaviour. The socio-cognitive framework informing teacher self-efficacy presents the construct as measurable, leaning towards a positivist epistemology, hence the quantitative bias reflected in the frequent use of Likert scale-based self-report questionnaires, producing numerical teacher self-efficacy ratings.

Current methods used to capture teacher self-efficacy present several limitations. First, the approach often results in losing the richness of the construct's complexity. As Tschannen-Moran and Woolfolk Hoy (2001) question, can single-item, or multiple-item, statements really capture the multifaceted dimensions of teacher self-efficacy? Furthermore, as Wheatley (2005, 749) observes, even though teacher self-efficacy is a continuous variable, due to the dominance of numerical scales, teachers are often presented 'in terms of two groups – those with "positive, high or greater" teacher self-efficacy, and those with "low, lower, or lesser" teacher self-efficacy". This labelling results in the illusion of a simple dichotomised construct, again overshadowing any possibility of complexity that might be present within the scales of self-efficacy. Furthermore, the broad labels are restrictive, offering a limited insight into potential types

of teacher self-efficacy (for example, subject, pedagogy, management) or the teachers' context.

Second, researchers are frequently concerned with instrument validity, reliability and specificity. For example, specificity concerns are evident in the production of multiple research instruments as outlined earlier. That said, generally speaking, research findings support the premise of self-efficacy (Pajares 1996). The main issue, however, concerning construct validity arises when researchers have attempted to develop instruments for measurement (Tschannen-Moran & Johnson 2011). One such issue is the extent to which a statement, or item, represents what it claims to represent. A second issue is finding an item's optimal level of specificity or external validity (Tschannen-Moran & Woolfolk Hoy 2001). That is, it is important to ensure that a measure's predictive power is not lost due to the specificity in skills and context being measured (rendering it useless for comparisons across teachers, classes, and subjects) nor that it is too general so that only broad comparisons can be made (Pajares 1996).

Finally, the validity of findings using Likert scale-based questionnaires assumes that a teacher's testimony is truthful. In contrast, research suggests that five-point Likert scales often sway respondents to restrict their preferences, avoiding extreme states such as strongly agree and strongly disagree (Cohen, Manion, and Morrison 2007), leading Dellinger, Bobbett, Olivier and Ellett (2008), for example, to reduce their scale to four-points. More specific to this research context, studies frequently conducted in parallel with professional development programmes are at risk of 'pretend teacher efficacy', in which they demonstrate inflated efficacy during professional development (Wheatley 2005). The human tendency to present oneself in the best possible light is a well researched phenomenon referred to as 'social desirability bias' (Maccoby & Maccoby 1954).

Researchers, including Labone (2004), Wheatley (2005) and Klassen et al. (2011) have called for an increase in qualitative methods in teacher self-efficacy, arguing that such approaches will offer a deeper understanding of how teacher beliefs and teacher self-efficacy operate. The few teacher self-efficacy studies that use qualitative approaches utilise wide-ranging methods. These include school observations (Ross & Bruce 2007), written reflections (Brand & Wilkins 2007), individual interviews (Cantrell & Callaway 2008), 'talk-aloud' protocols (Gabriele & Joram 2007) and open-ended questions in

questionnaires (Onafowora 2005). Often singular methods were used in studies with little attempt at triangulation. However, where multiple data sources were used in an attempt to triangulate the findings, the method of analysis, or analytical framework, were not made explicit, (for example, Ross & Bruce 2007; Hong 2010). Finally, most teacher self-efficacy research using qualitative approaches has focused on identifying sources of teacher self-efficacy (for example, mastery and vicarious experiences) rather than teacher self-efficacy *per se.*

In a response to the lack of triangulation concerning the explicit qualitative identification of self-efficacy, Wyatt (2013) detailed and developed a longitudinal case study exemplifying how data from several sources - observations and interviews - might collectively inform self-efficacy judgement. Such methods, however, require refinement as on occasion Wyatt interpreted interview responses concerning past successes, rather than future beliefs, as evidence for teacher self-efficacy.

To date, compared to quantitative studies, standardised qualitative self-efficacy method(s) or analytical framework(s) are yet to be developed or utilised widely. An exception is the recent study by Klassen and Durksen (2014) who developed a coding framework derived from extant theoretical and empirical research. The authors use quantitative methods, for example Bandura's (2006) work on scale construction to inform their coding. They also code responses for attributes associated with high self-efficacy (for example, increased effort and persistence) or low self-efficacy (for example, attributions to lack of ability, lack of agency).

Building on the methods of Klassen and Durksen (2014) and Wyatt (2013), this paper sought to explore and present an emerging framework utilising qualitative methods. Reflecting on the number of issues highlighted above concerning quantitative methods, the ability to widen the analytical perspective seemed essential.

Methods

This study was an interpretive multiple-case study conducted from a social constructivism perspective (Ball 2004; Denzin & Lincoln 2008). Reflecting Vygotsky's (1978) perspective of the theory of social constructivism, language in this study was considered an important mediator for participant and researcher meaning making (Blaikie 2007).

Study Context

We conducted the study alongside a two-year professional development programme – 'Thinking beyond the classroom' with the aim to enhance in-service secondary science teachers' pedagogy outside the classroom and co-construct ten outdoor science activities. The programme's activities and professional development were underpinned by elements of two pedagogical approaches with evidence for enhancing student attainment – Cognitive Acceleration through Science Education (CASE) (Shayer and Adey 2002) and Assessment for Learning (AfL) (Black & Wiliam 1998). See Glackin (2016) for an extended account of the professional development guiding principles, the content, the design of the professional development programme and the research design. The programme team included of one of the paper's authors (Glackin) who was a programme tutor, programme developer, as well as a researcher.

Participants

Eighteen secondary science teachers from ten secondary schools across Greater London enrolled onto the 'Thinking Beyond the Classroom' programme. During the first year of the programme, research data was collected with 12 of the 18 participating teachers. Five of the 12 participants withdrew from the programme and as such were not included in the study. Reasons for premature departure included: leaving the school (2), receiving promotion and having less time, feeling unable to continue without a colleague on the programme and feeling over committed to several professional development programmes. Furthermore, data from one participant was excluded from the study because they had not participated in the Year 1 observations or interviews.

The resulting participants became the six case studies teachers. Denzin and Lincoln (2008) define this group as a purposive sample – presenting the processes being studied – whereby the processes in this study included: teaching science outside and engagement in a professional development programme. Confidentiality was assured through the use of pseudonyms (British Education Research Association 2011). Stake (2005, 444) defines the case study as a 'bounded system', with Berg (2009) suggesting it is concerned with a particular person, social event or group. For this study, the bounded case study was the individual teacher who had completed the 'Thinking Beyond the Classroom' professional development (PD) programme. The programme was the study's backdrop (Bryman 2004) and the bounded time was almost two years. Table 1 summarises the case study teachers'

specialist science subject, general teaching experience, school type, accessibility to outdoor space and previous outdoor teaching experience.

-Table 1 insert here -

Research methods

Table 2 summarises the data sources collected over the two-year programme. Described briefly below are four methods (session questionnaires, a STEBI questionnaire, lesson observations and interviews) that were integral to teacher self-efficacy analysis.

-Table 2 insert here

Session questionnaires

Following each of the six professional development sessions teacher participants completed a questionnaire. The questionnaires asked teachers to rank their 'confidence' to teach each of the activities introduced during the session: For example, 'On a scale of 0-9 (0 being lowest), how confident do you feel to teach X?' (each activity would then be named). In addition to being invited to explain these ratings, teachers were asked to complete several open-ended questions concerning their future development/teaching. For example, 'reflect on one idea that you have been struck by today. How will you implement this in your classroom?'

The term confidence was used as a proxy for self-efficacy, as teachers expressed not understanding self-efficacy, saying it felt like 'jargon'. Acknowledging that confidence is a 'nondescript term that refers to strength of belief but does not necessarily specify what the certainty is about' (Bandura 1997, 382), we felt that by asking about particular and specific activities to be taught in the future, we were justified in deeming that responses reflected self-efficacy. Furthermore the open-ended questions offered teachers an opportunity for clarification and extended comments (Cohen, Manion, & Morrison 2007). In the findings we will therefore refer to the data as a self-efficacy.

STEBI questionnaire

The Riggs and Enoch's (1990) STEBI instrument, including Bleicher's (2004) recommendations for improved validity and reliability, were presented to colleagues at a research seminar at King's College London for comment and feedback. As the

development of a new STEBI reliability assessment was beyond the scope of the study, the discussion sought to find agreement on what further changes were permissible without the need for tool reassessment to narrow the specificity to science teaching outdoors. The concern was that the majority of previous research had used the questionnaire with in-service primary teachers, who are often considered to have low teacher efficacy in science teaching (Palmer 2006). As a result, the items were concerned with general science teaching and were not concerned with specific aspects of pedagogy that the 'Thinking Beyond the Classroom' programme focused on, for example, the use of practical work, teacher questioning, group work, or the use of the outdoors. Thus, in the seminar, items that easily lent themselves to slight word changes were discussed and agreed. The modified STEBI can be seen in Appendix 1. The items changed are listed in Appendix 3. The final STEBI questionnaire contained 12 items related to perceived self-efficacy and 12 items related to outcome expectancy.

Teachers completed the modified STEBI questionnaire, at the end of the programme. All participants were given a paper copy (with a stamped addressed envelope) and sent an electronic copy via email. Five case study teachers opted to complete and return the questionnaire electronically and one teacher returned it by post.

Lesson observations

A lesson observation tool was developed based on the pedagogical framework underpinning the programme and activities. The overarching aim for the lesson observations was to document teachers' interpretation of the professional development programme – including the use of the outdoors – and to substantiate what teachers reported during interviews. However, there were four discrete aspects of the activity that had specific pedagogical foci. These were: setting the scene, observing and collecting data, sharing and challenging ideas and linking ideas together/reflection. The observation tool provided observation prompts for particular aspects of the teaching. For example, observation prompts for the sharing and challenging ideas aspect included: *is the teacher eliciting/gathering students' ideas? Is the teacher encouraging students to challenge each other's ideas?*

Interviews

Semi-structured interviews with teachers took place following observed lessons with the interview duration ranging between 20-45 minutes. Given the research foci arose from a

constructivist epistemology, the interviews were regarded as an opportunity for the learners – in this case the teachers – to discuss their interpretation of the programme, and, more specifically, teaching science outdoors. Teachers' responses were considered to be a result of their personal belief system (Jones & Carter 2007), enabling an opportunity to elicit both explicit and more nuanced expressions of teacher self-efficacy. Interview questions included: *What determined the choice of the activity that you have trialed? Have you used the outdoors in lessons other than those from the programme?*

Analysis

Traditional quantitative analysis

We initially elicited teacher self-efficacy from the two quantitative sources: self-efficacy rating to teach programme activities (session questionnaires) and the STEBI questionnaires.

The case study teachers' self-efficacy ratings to teach programme activities were compiled and average ratings were calculated for individual case study teachers and individual activities (see Table 3).

Table 3 – in text here

Following Riggs and Enochs (1990) procedure, STEBI questionnaire scores were calculated for each case study teacher. As discussed above the item scores for each dimension – perceived self-efficacy and outcome expectancy – were aggregated for individual case study teachers. Table 4 presents individual teachers' STEBI scores, by dimension and overall.

Table 4 – here

Qualitative analysis

The qualitative analysis of self-efficacy was completed in four stages. The initial analysis informed an emerging analytical framework. The first stage was the identification of attributes or pedagogical practices for teachers with particular self-efficacy descriptions as highlighted in previous research findings. In the Literature Review, we listed teacher characteristics associated with high teacher self-efficacy. For example, high levels of planning, organization and enthusiasm (Allinder 1994; Muijs & Reynolds 2002) and a

willingness to experiment with new teaching methods and persist if initial implementation was flawed (Haney et al. 1996; Cousins & Walker 2000). The second stage was to organise and group the attributes and practices. As they arose from multiple research papers it became apparent that many attributes and practices described similar behaviours with different terminology. Appendix 2 lists studies that include a range of attributes and practices in their descriptions of teachers identified as having high self-efficacy. Emerging from these lists of attributes and practices were four dimensions: subject knowledge, flexibility/disposition, teacher/learner focus and behaviour management. Generally speaking, teachers with high self-efficacy would be expected to show high levels of all four dimensions. In order to think about the ways these dimensions would vary, behaviours for high and low teacher self-efficacy were described for each dimension. However, it was accepted that teachers might range between, and beyond, the framework's polarised descriptions.

The third stage of the qualitative analysis involved contextualising each dimension for the current professional development programme by developing descriptions of behaviour. As Bandura (1997) proposed, specificity of task is fundamental to teacher self-efficacy judgement; hence descriptions reflecting possible teachers' reactions to the 'Thinking Beyond the Classroom' programme were given. These descriptions consisted of teaching practice for teachers implementing the programme activities with high teacher self-efficacy. The development of the framework was iterative. That is, through the analysis of the six teacher case studies a more developed understanding of the four dimensions emerged which informed the revision of the contextual description. Table 5 presents the final framework.

Insert Table 5 here

The final stage of the qualitative analysis was the use of the framework. To this end, rich written descriptions (Becker 1970) were developed for each dimension for the case study teachers. Although the main data sources informing the descriptions were the interviews and lesson observations, all data was included (see Table 2). When conflicting evidence arose, further data was analysed. Where inconsistences remained, that is when teachers presented multiple and conflicting traits over the two years with no dominant behaviour, all attributes and practices were presented. 'Respondent validation' (Bryman 2004) was solicited from three case study teachers regarding the case study trustworthiness. The

updated case study descriptions were then compared to the framework and an interpretation of teacher self-efficacy given.

Findings

Overall, using quantitative approaches, teacher self-efficacy was high and little variation between the six case study teachers was seen (see Table 3 and 4). For example, the mean self-efficacy ratings to teach programme activities ranged from 6.2-7.2 – the mode was 7. Similarly, the STEBI data offered limited insight regarding teacher self-efficacy. All the case study teachers received a relativity high score on the perceived self-efficacy dimension - the range was 51-58 (from a possible score of 60). Case study teachers' outcome expectancy scores were not as high, ranging between 34-42 (from a possible score of 60). This finding of lower outcome expectancy scores was consistent with other studies (Riggs & Enochs 1990).

However, there was concern about the trustworthiness of the instruments as occasional disparities were evident between the reported rating and the interview as well as open-ended questionnaire data. For example, one of the case study teachers (Charlie) rated his self-efficacy to teach an activity *Materials All Around Us* activity as '7'; however in a later interview he admitted having felt initially under-confident following the session resulting in him not trialling the activity:

> *Interviewer: Why didn't you try Materials All Around Us before [today]?*
> *Charlie: I think because we didn't feel that confident about the lesson, and teaching the lesson, from the first PD session we had where we didn't really talk about how to teach it, we just talked about the tables and how we could change them. But after the follow-up PD session where we talked with you about it I think it became much clearer what the purpose of the lesson was. So I was quite happy to teach it after that. I think it's a really good lesson. I enjoyed it. I think the kids can get a lot out of it.*

In addition, three of the case study teachers, unprompted, annotated the STEBI questionnaire querying the exact meaning or context to which an item referred. For example, item 6 read: 'If students are underachieving in science, it is most likely due to ineffective science teaching'. Two teachers circled 'uncertain' (mid-point on the scale) with one teacher noting:

Some students will not engage whatever the strategy but generally I would probably agree.

Item 24 was another example where teachers expressed uncertainty: 'Even teachers with good science teaching abilities cannot help some kids learn science.' One teacher wrote:

In part this [item 24] may be due to the time required to turn the pupil around is simply not available when you are faced with a class of 30. On a one-to-one basis there is a greater chance of success.

The findings from the traditional analysis indicate that generally teachers had high self-efficacy to teach the science activities outside. However, inconsistences had emerged and the interpretation of some of the STEBI items was considered dependent on a teacher's context. To critique the traditional approaches further, and to present a qualitative method of analysis, below, for brevity and comparison, we present two case study teachers, Michael and Charlie. The case studies were chosen as they explicitly demonstrate different interpretations of self-efficacy dependent on the research method used. Such differences, although observed across the other cases, were more nuanced.

Case study teacher: Michael

Michael, an early career teacher, had a degree in physics. He taught general science at Key Stage 3 (KS3) (11-14 years) and Key Stage 4 (KS4) (14-16 years) and physics post-16 years. Prior to the programme his outdoor teaching consisted of sessions in which students measured speed and used field techniques. Michael had no previous formal training in Cognitive Acceleration through Science Education. He had, however, received considerable Assessment for Learning professional development.

Michael's quantitatively assessed teacher self-efficacy remained consistently high over the two-year programme. For example, Table 3 indicates that Michael gave 4 out of 9 activities ratings of 8 or above. Furthermore, Table 4 lists his STEBI score as 98 (out of a possible 120) – the highest score across the six case study teachers. In contrast, Michael's qualitatively assessed self-efficacy was described as moderate-low, evidence of which appeared across all dimensions, particularly visible in his lack of flexibility to adjust practice dominated by a teacher-centric approach. Michael did not integrate the programme's ideas into his practice, in total he taught outdoors on nine occasions, the

majority during the first year of the programme. He limited the activities he trialled to four out of a possible ten over the duration of the programme. Using our analytical dimensions, his self-efficacy is discussed.

Subject knowledge.

Despite being a teacher of general KS3 science, potentially due to his physics degree background, the activities Michael repeated most often were those most related to physics topics, for example, *Seeing the world through rose tinted glasses*, which encourages students to think about how they see colour. Furthermore, Michael's physics explanations were observed as more in-depth than his biology explanations. However, these explanations were often at the expense of opportunities for students to experience cognitive challenge. That is, Michael's teaching strategies often resulted in imparting science theory, rather than opening up and challenging thinking. Therefore, we begin to call into question the high self-efficacy found in traditional measures with respect to subject knowledge.

Flexibility and Disposition.

We considered Michael's disposition for teaching outside and for the professional development as apathetic – there was a sense of getting through activities and the programme sessions rather than engaging with them. Although he portrayed an understanding of pedagogical tools – such as Assessment for Learning – the strategies were often not evident in his practice. That said, Michael's attendance at five of six programme sessions and an attempt to trial some activities outside, suggested some willingness to experiment. However pedagogical risk taking was restrained, with limited change in pedagogical practice observed, and during the second year only two of the activities were taught with no repetition of activities trialled during Year 1. As a result, within the Flexibility and Disposition dimension Michael's behaviour was judged to reflect low teacher self-efficacy.

Teacher/learner focus.

Michael's focus was on his teaching rather than the students' learning. That is, during interviews he neglected to comment in any depth as to what particular groups/individual students might gain from different teaching approaches – other than that they will have fun. Michael described his role outside as a guide. This role of guide was observed during

lessons when he frequently led the whole class around the school grounds indicating relevant observations and scientific points of interest.

Summative testing occupied a prominent space in Michael's interviews and lessons, which we considered steered pedagogical decisions. This, combined with Michael's focus on being a guide, resulted in students often being told what they needed to know, with what appeared to be little regard for individual learners' needs. Further evidence of this predominantly teacher-centred approach was that Michael rarely used student names when addressing them directly, he seemed aloof and was occasionally critical of students' efforts during lessons.

Behaviour management.

Teachers identified as having high teacher self-efficacy implicitly incorporate behaviour management strategies into activities – feeling confident that students are able to manage their own learning. Michael, however, explicitly noted that he was 'not the master of behaviour management' and felt he should have control over students at all times when outside. This is in accordance with the teacher-centred approach described above. Furthermore, evident in the extract below was a resignation that students would always present challenging behaviour:

> *Well obviously just taking them out just has it's own challenges, different people wandering off.*

> *My role outdoors? [...] certainly managerial trying to keep them all, I mean that is like inevitable whenever you are teaching there seems like there is quite a large level of management involved.*

Michael's uncertainty of his ability and his resignation that there will be behaviour concerns suggested that he had low self-efficacy to manage student learning.

Case study teacher: Charlie

Charlie, like Michael, was an early career teacher. His first degree was biological sciences and he taught general science at KS3 and KS4 and biology post-16. Charlie's experience of outdoor teaching included organising a post-16 biology visit taught by an external field study provider and teaching KS3 ecology sampling skills in the school grounds. In

parallel with the outdoor professional development programme, Charlie also attended Cognitive Acceleration through Science Education professional development offered in-school.

Charlie's quantitatively assessed teacher self-efficacy was judged as moderate. Although his ratings to teach the activities were similar to Michael's (see Table 3), his STEBI score was joint lowest at 87 out of 120 (see Table 4). Furthermore, during the interviews he explicitly articulated concerns relating to his confidence when outside. For example, '[Inside] I feel more comfortable than outside' and 'We didn't feel that confident about the lesson, and teaching the lessons'.

Generally, the qualitative self-efficacy analysis paralleled the traditional quantitative self-efficacy analysis, however in Year 2 Charlie's self-efficacy using qualitative analysis was assessed as moderate-high. This increase in self-efficacy using the qualitative framework rested on the dimensions 'subject knowledge' and 'flexibility and disposition' when Charlie was observed as increasingly able to adapt his pedagogical practice to the outdoor environment. However, his view that students would misbehave outside, more than inside, and his unsubstantiated belief that he would not be able to manage such disruption suppressed his overall teacher self-efficacy. Compared to Michael (9), Charlie taught outside on more occasions (14), trialling six different activities, over the two-year programme. As with Michael, we discuss Charlie's teacher self-efficacy using the analytical dimensions.

Subject knowledge.

Charlie was observed as having good subject knowledge, enabling him to trial activities from across the sciences (1 biology-related, 1 chemistry-related, 2 physics-related and 2 scientific skill-related). His self-efficacy appeared to increase as his subject knowledge (and pedagogical practice) developed following the initial year of the programme. This increase was evident in: the integration of in situ outdoor examples into his explanations (especially in teaching physics related activities); increased ability to pose challenges pitched at the appropriate level; the development and use of effective open questions whilst outside; and the re-trialling of activities initially rejected in Year 1 – yielding positive outcomes. Hence, in the subject knowledge dimension Charlie displayed traits that suggested moderate-high self-efficacy.

Flexibility and Disposition.

Charlie was enthusiastic and open to ideas. On commencing the programme Charlie seemed quite flexible and open to change, trialling new aspects of pedagogy such as group work. However, apparent during Year 2 was an increase in flexibility, for example he was observed adapting and developing activities from the original plans for his specific context and class's ability. He also adapted the KS3 activities to use with his KS4 science classes. In addition, compared to Year 1, extended periods were spent outside, with less time exclusively used for student data collection, and more used for focused whole-class discussions around outdoor objects/specimens. Within this dimension Charlie demonstrated moderate-high teacher self-efficacy traits.

Teacher/student focus.

The changing teaching context influenced Charlie's teaching approach. That is, inside the classroom Charlie's approach was predominantly student-centred compared to when outside, where it was observed as teacher-centred. For example, inside activities were predominantly organised so that students had to work on tasks together; however, this was less evident outside, where students instead more often responded to teacher questions or collected data. As noted above, there was evidence of a shift during Year 2 whereby Charlie's practice outside became more student-centred, reflecting those practices more familiar in the classroom. Charlie's initial tendency to be teacher-centred, followed by a gradual shift to student-centred practice, was judged to reflect moderate teacher self-efficacy traits.

Behaviour management.

We viewed Charlie's student management skills as very good. However, similar to the former dimension, Charlie's approach to behaviour management was influenced by the context. Inside the classroom pedagogical choice dictated student behaviour, whereby the use of explicit behaviour management strategies were rarely observed. Conversely, outside, the management of students often dictated the pedagogical approach chosen. This practice seemed related to a 'fear' that students would misbehave:

> *I'm more comfortable inside but it would help if we did do it outside I think as the examples are around us easily and we can look at them again etc. I use the board a lot to focus their ideas, and that isn't there. I think it is more me, and trusting them. Trying to gather their ideas might be difficult, they might be difficult; there*

might be too many distractions. But saying that they might get over this with practice.

Related to this 'fear' was a concern during the interviews that colleagues might observe his students behaving poorly. Charlie reported only positive behaviour of students whilst outside, and although some change was noted in Charlie's teaching approach, he continued to articulate the fear of potential student misbehaviour when outside. Hence, Charlie's uncertainty concerning his ability to manage students outside suggested low teacher self-efficacy.

Summary

By using the qualitative framework alongside traditional quantitative self-efficacy analysis, a rich representation was captured of how the case study teacher felt about their skills to teach outside. So whilst Michael's quantitatively assessed teacher self-efficacy was 'high' and Charlie's was 'moderate', their self-efficacy assessed using qualitative methods at the end of the programme were summarised as 'moderate-low' and 'moderate-high', respectively. Hence, two advantages of the analytical framework have emerged. First, the framework offered a contrasting outcome to that presented using traditional methods and has justified questions concerning social desirability bias. Second, the analysing framework offered a more nuanced insight of self-efficacy so slight differences between articulated self-efficacy and embodied self-efficacy might be recognised.

Discussion

In this article we sought to explore how a qualitative approach might be used to gain a more comprehensive understanding of teachers' self-efficacy and whether such approaches deserve further development and future consideration. To address the first research question: *How might qualitative methods be used to gain a more comprehensive picture of teacher self-efficacy*? we have presented a qualitative analytical framework for self-efficacy informed by attributes and practices identified in previous teacher self-efficacy research. The four dimensions, along with the high/low self-efficacy trait statements and context exemplification enabled a wide range of data sources to be analysed, compared and contrasted to create each teacher case study. In developing this approach we are acknowledging the complexity of the construct. That is, first, the multiple dimensions of self-efficacy need to be identified; second, the interrelationship of

self-efficacy dimensions needs consideration, as does third, the influence of the specific teaching context.

Several researchers have called for the development of qualitative approaches to teacher self-efficacy, for example, questioning instrument validity (Tschannen-Moran & Woolfolk Hoy 2001) or context specificity (Pajares 1996). However, there are no explicit studies that compare the approaches used to identify self-efficacy that might highlight the limitations of the sole use of traditional quantitative methods. Hence, the second research question this article sought to explore was: *What, if any, variation is there between teacher self-efficacy judgement when qualitative and quantitative data methods are applied?* We addressed the question by collecting multiple data sources that enabled both a traditional quantitative approach, for example, self-efficacy ratings and STEBI scores, alongside the use of the qualitative analytical framework.

By using the two approaches we found variation in self-efficacy judgements that we presented in two teacher case studies. For Michael, initially using the quantitative traditional approach a picture emerged of high teacher self-efficacy and low programme implementation. Without the qualitative analysis, his lack of implementation or integration of programme activities might be understood in two ways. It could be that Michael did have high self-efficacy and just did not agree with the way the professional development programme was provided. His philosophy towards teaching might have been completely distinct to that of the programme. And yet, he attended most of the training days in the programme. He could have withdrawn from either the programme or the research or both, as did some of the other attendees. That he carried on suggested he was not at odds with the philosophy to the extent that he showed a lack of self-efficacy in implementing strategies from the programme. Alternatively, Michael's high self-efficacy might have given him the tools to carry on teaching in a way that he thought was effective without needing to change his practices. Southerland, Sowell, Blanchard and Granger (2011) have termed a similar condition, when a teacher has high self-efficacy but low programme implementation, as 'pedagogical contentment'. They suggest the omission of professional doubt, possibly caused by a lack of reflection and evaluation, leads to the lack of engagement and change in practice. That is, Michael's high self-efficacy need not be questioned.

On the other hand, when the qualitative analytical framework is introduced and a richer picture offers moderate-low self-efficacy assessment, a third interpretation for Michael's lack of programme implementation emerges. In this case, he may present a façade that shows social desirability bias, and protects his sense of self: that he is a good teacher and can manage the tasks of teaching in a way that is competent. That is, like many professionals, teachers are often wary of presenting flawed/sub-standard qualities that might damage their professional status. Further, Michael had invested significant time into the programme and might have sensed he was supposed to feel confidence. Also, chiming with Wheatley's (2005) findings, there was a risk of 'pretend teacher efficacy' as the multiplicity of the educational researcher's role, as tutor and programme developer might have compromised his response. Hence, what appears in a deeper investigation is the complexity around Michael's sense of self. That is, he came across as confident, he hesitated to attempt changes to his practice due to an underlying doubt about whether the outcomes would be approved.

Clearly, any of the three interpretations may be correct. It is even possible that each holds a portion of the truth for Michael at different times. Our point here is that when traditional quantitative approaches are triangulated and integrated with qualitative approaches a more complete picture emerges offering a more dynamic and complex insight of a teacher's self.

Compared to Michael, Charlie had moderate teacher self-efficacy but was considered to have moderate-high self-efficacy when judged using the qualitative framework at the end of Year 2. Again, just relying on quantitative measures of self-efficacy as provided by Charlie himself, we might predict that his experience with the programme would provide a similar outcome to that of Michael. And yet, use of other sources and analysis suggests otherwise. Over the course of the professional development programme Charlie's self-efficacy increases in a way that Michael's does not. The qualitative method of assessing Charlie's judgement of his ability enabled a comparison between data sources offering a richer, as well as a more nuanced, insight into Charlie's development over the course of the programme.

Michael and Charlie's case studies were discussed in this paper as they explicitly demonstrate how self-efficacy outcomes are dependent on the measure used. However, Table 4 presents that 2 out of 6 teachers have the same self-efficacy outcome ('high' on

both measures) and that there was no variation between teacher self-efficacy judgement when qualitative and quantitative data methods were applied. Although broadly the self-efficacy outcomes were considered consistent, what the self-efficacy framework offered when applied to these two case study teachers was an opportunity to explore specific aspects of self-efficacy and tease out, however slight, emerging differences. Hence, rather than a redundant measure for these case studies, the framework offered a more nuanced insight of their self-efficacy.

Limitations and Future Research

The development and use of qualitative approaches is not straightforward or free from difficulties. The use of a qualitative analytical framework is time-consuming and possibly costly. One reason the quantitative measures of teacher self-efficacy are so ubiquitous stems from the ease with which they can be used. We argue, evidenced in this study, that the benefits of added credibility outweigh the costs of time needed to use qualitative approaches. However, we acknowledge that the framework is only emerging and requires further development, as do the ways of using it, before it can be more widely taken up within the field.

In addition, caution is required when aligning teaching practice/traits with specific levels of teacher self-efficacy. It is possible that a range of pedagogical practices could result in a teacher feeling different levels of judgement towards their future teaching, resulting in an accumulative 'general' self-efficacy. That said, different aspects have been identified as more or less influential on 'general' teacher self-efficacy. For example, Mulholland and Wallace (2001) found that whilst subject knowledge influenced science teachers' self-efficacy and resulting practice, teacher self-efficacy for managing student learning was equally important. The researchers noted that teachers with high teacher self-efficacy for subject knowledge but low teacher self-efficacy for managing student learning avoided hands-on science activities and teacher demonstrations often replaced whole-class inquiry activities (Mulholland & Wallace 2001). Whilst on one hand this finding, that multiple states of teacher self-efficacy act to influence pedagogical choice, highlights just one of the challenges of understanding teacher self-efficacy, on the other hand it illuminates the importance of a descriptive complimentary analytical framework so that specific judgements might be better understood.

Finally, further research is required to refine, adjust and adapt the qualitative analytical framework so that it can be used within other contexts. The approach and framework are in their infancy. For example, terms used within the framework might need to be developed so that they are more generic and transferable to other settings. Whilst we support Wheatley's (2005) concern that the current use of labelling gives the illusion of a simple dichotomised construct, as an intermediary step for qualitative methods to develop, to allow greater transferability and to act as a bridge between approaches, we propose that more details of the four dimensions for very high, very low and moderate teacher self-efficacy are required. However, we are currently not satisfied with the use of such labels and would in future develop a different summary method. Using the qualitative framework across a number of different research contexts should enable agreed terms to be developed and tested.

Implications for professional development

Teacher self-efficacy measures are frequently used for determining the success of professional development programmes. The outcome usually informs decisions concerning future implementation and, to a lesser extent, programme design. Through the development and use of broader approaches to self-efficacy assessment, we suggest a more accurate insight will be gained of a programme's success in developing teachers' judgement in their ability relating to a particular pedagogical practice. Further, the rich descriptions required for the four dimensions offer a depth of insight of the interrelationship of teacher's judgement of their ability in a particular context unattainable through quantitative methods. Such rich descriptions, particularly if constructed periodically during a programme, offer knowledge of how specific self-efficacy dimensions change (for example, subject knowledge and behaviour management). Hence, this information could influence programme strategies and foci: If some teachers react with hesitancy, it might be that they require a different approach to help them implement what is considered to be good practice in teaching. This outcome would substantiate research concerning sources of teacher self-efficacy related to professional development (Tschannen-Moran & Johnson 2011; Palmer 2011).

Conclusion

Ensuring that teachers develop high teacher self-efficacy is important for the success of the teaching profession (Bandura 1997). Although judgement of self-efficacy is important to researchers, policy makers and teacher educators, the assessment has been dominated

by traditional quantitative approaches (Klassen et al. 2011), to date offering a limited understanding of the concept. By exploring the self-efficacy of two secondary science teachers participating in a two-year professional development programme using both traditional quantitative scales and qualitative methods including interviews and lesson observations, the findings suggest that traditional quantitative scales do not fully capture teacher self-efficacy and that there are inconsistencies between self-efficacy assessment outcomes when different research approaches are used. We have argued that to achieve a more complete and comprehensive picture of teacher self-efficacy, it is essential that traditional quantitative approaches are better triangulated and integrated with other sources of data. We have offered an emerging qualitative analytical framework and invite further developments on the framework and for use within other contexts.

References

Allinder, R.M. 1994. "The relationship between efficacy and the instructional practices of special educational teachers and consultants". *Teacher Education and Special Education* 17: 86-95. doi: 10.1177/088840649401700203.

Armor, D., P. Conroy-Oseguera, M. Cox, N. King, L. McDonnell, A. Pascal, E. Pauly, and G. Zellman. 1976. *Analysis of the school preferred reading programs in selected Los Angeles minority schools*. Santa Monica, CA: RAND (ERIC Document Reproduction Service No. ED 130 243).

Ashton, P., D. Buhr, and L. Crocker. 1984. "Teachers' sense of efficacy: A self- or norm referenced construct." *Florida Journal of Educational Research* 26 (1):29-41.

Ashton, P.T., and R.B. Webb. 1986. *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman Press.

Ball, S.J., ed. 2004. *The routledgefalmer reader in sociology of education*. London: RoutledgeFalmer.

Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Bandura, A. 2006. "Guide for constructing self-efficacy scales." In *Self-efficacy beliefs of adolescents*, edited by F. Pajares and T. Urdan. Greenwich, CT: Information Age Publishing.

Becker, H. 1970. *Sociological work: Method and substance*. Chicago: Aldine.

Berg, B. L. 2009. *Qualitative research methods for the social sciences*. Boston: Pearson.

Black, P., and D. Wiliam. 1998. *Inside the black box*. London: nferNelson.

Blaikie, N. 2007. *Approaches to social enquiry*. Second ed. Cambridge: Polity.

Bleicher, R. 2004. "Revisiting the STEBI-B: Measuring self-efficacy in preservice elementary teachers." *School Science and Mathematics* 104 (8):383 – 391. doi: 10.1111/j.1949-8594.2004.tb18004.x.

Brand, B.R., and J.L.M. Wilkins. 2007. "Using self-efficacy as a construct for evaluating science and mathematics methods courses." *Journal of Science Teacher Education* 18 (2):297-317. doi: 10.1007/s10972-007-9038-7.

British Education Research Association. 2011. "Ethical guidelines for educational research." Last Modified March 2016. https://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-Ethical-Guidelines-2011.pdf?noredirect=1.

Bryman, A. 2004. *Social research methods*. Second ed. Oxford: Oxford University Press.

Cantrell, P., S. Young, and A. Moore. 2003. "Factors affecting science teaching efficacy of preservice elementary teachers." *Journal of Science Teacher Education* 14:177-192. doi: 10.1023/A:1025974417256.

Cantrell, S.C., and P. Callaway. 2008. "High and low implementers of content literacy instruction: Portraits of teacher efficacy." *Teaching and Teacher Education* 24 (7):1739-1750. doi: /10.1016/j.tate.2008.02.020.

Chacon, C.T. 2005. "Teachers' perceived efficacy among English as a foreign language teachers in middle school in Venezuela." *Teaching and Teacher Education* 21:257-272. doi: /10.1016/j.tate.2005.01.001.

Chesnut, S. R., and H. Burley. 2015. "Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis." *Educational Research Review* 15:1-16. doi: http://dx.doi.org/10.1016/j.edurev.2015.02.001.

Cohen, L., L. Manion, and K. Morrison. 2007. *Research methods in education*. London: Routledge.

Cousins, J.B., and C.A. Walker. 2000. "Predictors of educators' valuing of systemic inquiry in schools." *Canadian Journal of Program Evaluation* (Special Issue): 25-53.

Czerniak, C. M., and M. L. Schriver. 1994. "An examination of preservice science teachers' beliefs and behaviors as related to self-efficacy." *Journal of Science Teacher Education* 5 (3):77-86. doi: 10.1007/BF02614577.

De Laat, J., and J. Watters. 1995. "Science teaching self-efficacy in a primary school: A case study" *Research in Science Education* 25: 4553-4464. doi: 10.1007/BF02357387.

Dellinger, A., J. Bobbett, D.F. Olivier, and C.D. Ellett. 2008. "Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self." *Teaching and Teacher Education* 24 (3):751-766. doi: /10.1016/j.tate.2007.02.010.

Denzin, N.K., and Y.S. Lincoln. 2008. "Introduction: The discipline and practice of qualitative research." In *The landscape of qualitative research* edited by N.K. Denzin and Y.S. Lincoln, 1-44. London: Sage.

Gabriele, A. J., and E. Joram. 2007. "Teachers' reflections on their reform-based teaching in mathematics: Implications for the development of teacher self-efficacy." *Action in Teacher Education* 29 (3):60-74. DOI: 10.1080/01626620.2007.10463461

Gardner, P. L. 1996. "The dimensionality of attitude scales: A widely misunderstood idea." *International Journal of Science Education* 18 (8):913–919. doi: 10.1080/0950069960180804.

Gibson, S., and M. H. Dembo. 1984. "Teacher efficacy: A construct validation." *Journal of Educational Psychology* 76 (4):569-582. doi: 10.1037/0022-0663.76.4.569.

Glackin, M. 2016. "'Risky fun' or 'authentic science'? How teachers' beliefs influence their practice during a professional development programme on outdoor learning." *International Journal of Science Education* 38 (3):409-433. doi: http://dx.doi.org/10.1080/09500693.2016.1145368.

Goddard, R.D., W.K. Hoy, and A. Woolfolk Hoy. 2004. "Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions." *Educational Researcher* 33 (3):3-13. doi: 10.3102/0013189X033003003.

Haney, J.J., C.M. Czerniak, and A.T. Lumpe. 1996. "Teacher beliefs and intentions regarding the implementation of science education reform strands." *Journal of Research in Science Teaching* 33 (9):971-993. doi: 10.1002/(SICI)1098-2736(199611)33:9<971::AID-TEA2>3.0.CO;2-S.

Henson, R.K. 2002. "From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research." *Educational Psychologist* 37 (3):137-150. doi: 10.1207/S15326985EP3703_1.

Hong, Ji Y. 2010. "Pre-service and beginning teachers' professional identity and its relation to dropping out of the profession." *Teaching and Teacher Education* 26 (8):1530-1543. doi: http://dx.doi.org/10.1016/j.tate.2010.06.003.

Jones, M.G., and G. Carter. 2007. "Science teacher attitudes and beliefs." In *Handbook of research on science education*, edited by S. Abel and N. Lederman. London: Lawrence Erlbaum Association.

Klassen, R. M., and T. L. Durksen. 2014. "Weekly self-efficacy and work stress during the teaching practicum: A mixed methods study." *Learning and Instruction* 33:158-169. doi: http://dx.doi.org/10.1016/j.learninstruc.2014.05.003.

Klassen, R.M., V. Tze, S.M. Betts, and K.A. Gordon. 2011. "Teacher Efficacy Research 1998–2009: Signs of Progress or Unfulfilled Promise?" *Educational Psychology Review* 23 (1):21-43. doi: 10.1007/s10648-010-9141-8.

Labone, E. 2004. "Teacher efficacy: maturing the construct through research in alternative paradigms." *Teaching and Teacher Education* 20: 341-359. doi: 10.1016/j.tate.2004.02.013.

Maccoby, E. E., and N. Maccoby. 1954. "The interview: A tool of social science." *Handbook of social psychology* 1:449-487.

Muijs, D., and D. Reynolds. 2002. "Teachers' beliefs and behaviors: What really matters." *Journal of Classroom Interaction* 37:3-15.

Mulholland, J., and J. Wallace. 2001. "Teacher induction and elementary science teaching: Enhancing self-efficacy." *Teaching and Teacher Education* 17 (2):243-261. doi: 10.1016/S0742-051X(00)00054-8.

Onafowora, L.L. 2005. "Teacher efficacy issues in the practice of novice teachers." *Educational Research Quarterly* 28 (4):9.

Pajares, F. 1996. "Self efficacy beliefs in academic settings." *Review of Educational Research* 66 (4):533-578. doi: 10.3102/00346543066004543.

Palmer, D. H. 2006. "Sources of self-efficacy in a science methods course for primary teacher education students." *Research in Science Education* 36 (4):337-353. doi: 10.1007/s11165-005-9007-0.

Palmer, D. H. 2011. "Sources of efficacy information in an inservice program for elementary teachers." *Science Education* 95 (4):577-600. doi: 10.1002/sce.20434.

Pruski, L.A., S.L. Blanco, R.A. Riggs, K.K. Grimes, C.W. Fordtran, G.M. Barbola, J.E. Cornell, and M.J. Lichtenstein. 2013. "Construct Validation of the Self-Efficacy Teaching and Knowledge Instrument for Science Teachers-Revised (SETAKIST-R): Lessons Learned." *Journal of Science Teacher Education*:1-24. doi: 10.1007/s10972-013-9351-2.

Riggs, I. 1995. "The characteristics of high and low efficacy elementary teachers." The annual meeting of the National Assocation of Research in Science Teaching San Francisco, CA, April 1995.

Riggs, I. M., and L. G. Enochs. 1990. "Toward the development of an elementary teacher's science teaching efficacy belief instrument." *Science Education* 74 (6):625-637. doi: 10.1002/sce.3730740605.

Ross, J. A. (1998). "The antecedents and consequences of teacher efficacy". In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 7). Greenwich, CT: JAI Press.

Ross, J.A., and C.D. Bruce. 2007. "Teacher self-assessment: A mechanism for facilitating professional growth." *Teaching and Teacher Education* 23 (2):146-159. doi: 10.1016/j.tate.2006.04.035.

Rubeck, M.L., and L.G. Enochs. 1991. "A path analytical model of variables that influence science and chemistry teaching self-efficacy and outcome expectance in

middle school science teachers." The annual meeting of the National Association for Research in Science Teaching, Fontana, WI, April.

Shayer, M. , and P. Adey. 2002. *Learning Intelligence: Cognitive Acceleration across the curriculum from 5 to 15 years*. Buckingham: Open University Press.

Southerland, S. A., S. Sowell, M. Blanchard, and E. Granger. 2011. "Exploring the construct of pedagogical discontentment: A tool to understand science teachers' openness to reform." *Research in Science Education* 41 (3):299-317. doi: 10.1007/s11165-010-9166-5.

Stake, R.E. 2005. "Qualitative case studies." In *The Sage handbook of qualitative research*, edited by N.K. Denzin and Y. S. Lincoln. London: Sage.

Tschannen-Moran, M., and D. Johnson. 2011. "Exploring literacy teachers' self-efficacy beliefs: Potential sources at play." *Teaching and Teacher Education* 27:751-761. doi: 10.1016/j.tate.2010.12.005.

Tschannen-Moran, M., and A.W. Woolfolk Hoy. 2001. "Teacher efficacy: Capturing an elusive construct." *Teaching and Teacher Education* 17 (7):783-805. doi: 10.1016/S0742-051X(01)00036-1.

Vygotsky, L.S. 1978. *Mind in society: The development of the higher psychological processes*. Cambridge, MA: Harvard University Press.

Wheatley, K.F. 2002. "The potential benefits of teacher efficacy doubts for educational reform." *Teaching and Teacher Education* 18 (1):5-22. doi: 10.1016/S0742-051X(01)00047-6.

Wheatley, K.F. 2005. "The case for reconceptualizing teacher efficacy research." *Teaching and Teacher Education* 21 (7):747-766. doi: 10.1016/j.tate.2005.05.009.

Woolfolk, A.E., B. Rosoff, and W.K. Hoy. 1990. "Teachers' sense of efficacy and their beliefs about managing students." *Teaching and Teachers Education* 6:137-148. doi: 10.1016/0742-051X(90)90031-Y.

Wyatt, M. 2012. "Towards a re-conceptualization of teachers' self-efficacy beliefs: tackling enduring problems with the quantitative research and moving on." *International Journal of Research & Method in Education* 59 (2):217-242. doi: 10.1080/1743727X.2012.742050.

Wyatt, M. 2013. "Overcoming low self-efficacy beliefs in teaching English to young learners." *International Journal of Qualitative Studies in Education* 26 (2):238-255. doi: 10.1080/09518398.2011.605082.